

# QUALITY CONTROL DEVICE FOR VOICE PACKET COMMUNICATIONS

## BACKGROUND OF THE INVENTION

### 1. Field of the Invention

The present invention relates to a quality control device for voice packet communications that uses a packet network of, for example, the Internet.

### 2. Description of Related Art

Recently, techniques are been proposed for transmitting a voice signal in real time through a packet network of, for example, the Internet, and devices therefor are being introduced for actual use.

However, the Internet was originally developed for data communications that do not require real time transmission, and the packet transmission on the Internet is not guaranteed qualitatively. Therefore, there is the possibility that phenomena, such as packet-lacking (packet loss), delay, and jitter, that deteriorate a decoded voice will occur on the Internet.

Therefore, if the Internet is used for a communications function, such as telephone communications, that require real-time responsiveness, a buffer device is needed to prevent transmission intermittence.

Let us assume that this buffer device stores voice packets (note that these voice packets include encoded voice data compressed according to an irreversible compressing/encoding method in many cases), which have been

received from a network, for example, the Internet, in the order of reception, and reads them in the order of storage.

In this case, the reading is always and repeatedly carried out at intervals of a fixed decoding unit time that is required by a decoding circuit that decodes (decompresses) the encoded voice data.

Therefore, in a case where this buffer device is used, if the arrival of a voice packet at a receiver is delayed for more than the fixed time because of, for example, the influence of jitter, the voice packet is not stored onto the buffer device, and only the reading continues, and, as a result, voice packets to be read out will be exhausted.

Since there is a need to keep voice packets being supplied to the decoding circuit at the intervals of the decoding unit time even when such exhaustion occurs, a technique for inserting a complementary packet that contains predetermined voice data (in many cases, this is voice data that generates a slight noise near voice-absence as a decoded voice) is generally used in this case.

However, when the complementary packet is inserted, a packet whose arrival has been delayed by the above-mentioned jitter is sent later. Therefore, disadvantageously, the number of packets in the buffer device gradually increases, and a transmission delay is lengthened with the lapse of time.

If the transmission delay becomes long, the response to the contents of speech, for example, in a bidirectional

conversational voice will be unnaturally delayed, and the quality of communication will fall.

A possible countermeasure against this is to at first delete (discard) the voice packet that has been stored in the buffer device (i.e., voice packet that occupies the top position) when the number of stored voice packets exceeds a predetermined number.

Another possible countermeasure is that a position where the complementary packet is inserted is fixed at this top position when the arrival of a voice packet is delayed for more than a fixed time, and voice packets to be read out are exhausted.

However, if the complementary packet is inserted or the voice packet is deleted at only about the top position in this way, an advantage of being able to simplify the processing can be obtained, but, only the state of the top position (i.e., state of the voice packet that has been read out prior to that) can be monitored. Therefore, as a result, the possibility that such deletion and insertion will be successively carried out for a specific position on a series of voice packets increases.

If deleted, effective voice data needed when decoded will be lost, and, if inserted, unnecessary voice data will be mixed when decoded. Therefore, these are operations that deteriorate the quality of decoded voice output, and, if the deletion or insertion is successively carried out from or onto the series of voice packets, the possibility that